

(19)



Europäisches Patentamt
European Patent Office
Office européen des brevets



(11)

EP 0 487 229 B1

(12)

EUROPEAN PATENT SPECIFICATION

(45) Date of publication and mention
of the grant of the patent:
24.09.1997 Bulletin 1997/39

(51) Int Cl.⁶: **C12N 15/12, C07K 14/435**

(21) Application number: **91310334.7**

(22) Date of filing: **07.11.1991**

(54) **A DNA encoding a protein which binds to the enhancer of alpha-fetoprotein gene**

Für ein Protein, das sich mit dem Enhancer des Alpha-Fetoprotein verbindet, kodierende DNS

DNA codant pour une protéine se liant au gène de l'agent de renforcement de l'alpha fetoprotéine

(84) Designated Contracting States:
AT BE CH DE DK ES FR GB IT LI LU NL SE

(30) Priority: **07.11.1990 JP 301412/90**

(43) Date of publication of application:
27.05.1992 Bulletin 1992/22

(73) Proprietor: **SNOW BRAND MILK PRODUCTS CO., LTD.**
Sapporo-shi, Hokkaido 065 (JP)

(72) Inventors:
• **Morinaga, Tomonori 3A, Kurai Manshon, Tochigi-ken (JP)**
• **Yasuda, Naotaka 2A, Okada Manshon, Tochigi-Ken (JP)**
• **Higashio, Kanji Saltama-Ken (JP)**
• **Tamaoki, Taiki Calgary Alberta (CA)**

(74) Representative: **Davies, Jonathan Mark et al Reddie & Grose**
16 Theobalds Road
London WC1X 8PL (GB)

(56) References cited:
EP-A- 0 427 863

- **MOLECULAR AND CELLULAR BIOLOGY** vol. 8, no. 12, December 1988, WASHINGTON, DC USA pages 5179 - 5187; Sawadaishi, K. et al.: 'Interaction of a Hepatoma- Specific Nuclear factor with transcription-regulatory sequences of the HumanAlpha-Fetoprotein and albumin genes.'
- **BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS**. vol. 143, no. 1, 27 February 1987, DULUTH, MINNESOTA US Wang, Zhi; Cote, Gilbert J.; Chiu, Jen Fu: 'A specific .alpha.-fetoprotein gene binding protein in .alpha.-fetoprotein producing rat hepatomas'
- **JOURNAL OF BIOLOGICAL CHEMISTRY**. vol. 262, no. 10, 5 April 1987, BALTIMORE US pages 4812 - 4818; WATANABE, K.ET AL.: 'Cell-specific enhancer activity in a far upstream region of the human alpha fetoprotein gene'
- **Cell** (1989) 59:145

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

EP 0 487 229 B1

Description

The present invention concerns a DNA sequence encoding a protein which specifically binds to the enhancer of α -fetoprotein gene and enhances the transcription of α -fetoprotein gene.

Expression of the genetic information in eukaryotes is regulated at the steps of transcription, RNA processing, translation and processing after the translation. Of all the regulation steps, that at the transcription level gives the stronger effect on the expression of the genetic information.

In the 5'-flanking region of genes which are transcribed by RNA polymerase II a promoter and, in some cases, enhancers which are involved in the transcription of the gene are present.

Recently, nuclear factors which bind to promoters or enhancers by recognizing the specific nucleotide sequences in the promoter or the enhancer were found. And it was found that the activity of promoter was regulated by the binding of these nuclear factors. The proteins which enhance transcription by RNA polymerase by binding to TATA box, CAAT box or GC box have been studied and the cDNAs for these proteins have been cloned.

The studies on the factors that bind to enhancers are also in progress. For example, cDNAs of OTF-2 (Octamer transcription factor-2) [Michael et al., Nature 336, 544-551, (1988)], which is B-cell specific and binds to the enhancer of immunoglobulin κ -chain gene, and OTF-1, which recognizes the same nucleotide sequence and present ubiquitously in many tissues, have been isolated.

In the human α -fetoprotein gene, the presence of a region having enhancer activity at 3.5 kb upstream of the promoter was found.

A factor which binds to this enhancer was found and was termed AFP-1 [Sawadaishi et al., (Molecular and Cellular Biology), vol.8, 5179-5187(1988)]. However, the protein structure or the physicochemical characteristics of AFP-1 is unknown and the gene encoding AFP-1 has not been isolated.

In the investigation of the nuclear factor which specifically binds to the region having TTAATAATTA core sequence present in the enhancer of the human α -fetoprotein gene, the inventors have isolated a cDNA encoding the factor and determined the nucleotide sequence of the cDNA.

Frain et al, described a DNA-binding protein termed LF-B1 which they purified from rat liver nuclei [Cell 59, 145-157 (1989)]. LF-B1 binds AT-rich sequences in the promoters of many liver-specific genes and activates their expression. The complete nucleotide sequence of the cDNA for rat LF-B1 was determined and the amino acid sequence was deduced. Although LF-B1 binds the nucleotide sequences that resemble the target sequence for the DNA-binding protein described in the present application, these two proteins are distinct from

each other.

A portion of the cDNA for the enhancer-binding protein described in the present application has been described in a European patent application, EP A 0 473 863, unpublished at the priority date of this application.

The present invention offers a novel DNA sequence which codes for the protein that specifically binds to the enhancer of the human α -fetoprotein gene. This DNA may be useful for the production of biologically active proteins by inserting this DNA into an expression vector and co-transfecting animal cells with another expression vector in which a gene for the biologically active protein is placed under the control of α -fetoprotein gene enhancer and promoter.

The present invention concerns a DNA having a nucleotide sequence shown in Fig 8. and encoding a protein which binds specifically to the human α -fetoprotein gene enhancer. The nucleotide sequence shown in Fig 8. is an example of such DNA sequence. This invention includes such DNAs which contain point mutations from which replacement of amino acid residues could result. An example of such mutation is the mutation at nucleotide 7508 in the sequence shown in Fig. 8. This mutation resulted in a replaced amino acid residue.

Fig.1 shows restriction enzyme map of the enhancer region present in the 5'-flanking region of the human α -fetoprotein gene. The region from 3.3kb to 3.7kb upstream of the transcription initiation site is presented magnifying the scale. H and Ha represent recognition sites for restriction enzymes Hind III and Hae III, respectively.

Fig.2 shows construction of CAT plasmid pAF1.OE25 CAT. A plasmid pAF1.OE25CAT is constructed by inserting Nla IV/Xho I fragment (31bp) of the enhancer region into the Cla I site of plasmid pAF1.OCAT.

Fig.3 shows the result of a CAT assay using HuH-7 human hepatoma cells. As shown, Nla IV/Xho I fragment has an enhancer activity.

Fig.4 shows the procedure to clone cDNA synthesized from mRNA into λ gt11 expression vector. A vertical arrow in the figure indicates the site in the lacZ gene to where cDNA is inserted.

Fig.5 shows restriction map of the cDNA and positions of partial cDNA clones.

Fig.6 shows the results of gel mobility shift assay with extracts prepared from λ 2 and λ gt11 lysogens. Control indicates an assay without lysogen extract. Two shifted bands (a and b) appear only when extract prepared from λ 2 lysogen is used.

Fig.7 shows the result of DNase I footprint analysis with λ 2 lysogen extract. G + A represents a size marker which was prepared by the chemical cleavage of the probe DNA according to the method of Maxam and Gilbert. DNA extracted from the band which was not shifted was run in lane 1. Lanes 2 and 3 show DNA extracted from shifted bands a and b, respectively. Panel A shows the result obtained by labelling the 5'-terminus of Sma I

site. Panel B shows the pattern obtained by labelling the 5'-terminus of Hinc II site.

Fig.8 shows the nucleotide sequence of AFP-1 cDNA and amino acid sequence deduced from it. Nucleotides that differ between two overlapping clones are indicated by closed triangles (▼). Asterisk (*) indicates adenine missing in λ 488. Homeodomains are indicated by solid underlines and zinc-fingers are indicated by broken underlines.

The protein described in their invention binds to the enhancer of human α -fetoprotein gene and has a primary amino acid sequence shown in Fig.8.

One can prepare such DNA from mRNA extracted from the α -fetoprotein-producing cells of hepatocyte lineage. HuH7, HepG2 and Hep3B are the examples of such cell which produce α -fetoprotein.

In the present invention, cells producing α -fetoprotein were cultured and RNA was extracted from the cells. Poly(A) RNA was prepared from the RNA by means of oligo(dT) cellulose chromatography according to the standard procedure.

An example of the procedure to be used to extract RNA from cultured cells is guanidinium isothiocyanate-cesium chloride procedure [J.M. Chirgwin et al., (Biochemistry 18, 5294 (1979))].

cDNA was synthesized using poly(A) RNA collected as mentioned above as template for reverse transcriptase. After the formation of double stranded cDNA from the single stranded cDNA. The product was ligated to λ gt11 DNA and recombinant phage was constructed by packaging the recombinant DNA into λ -phage. Commercially available packaging system, packaging system, for example (Promega), can be used for λ -phage packaging.

Then, recombinant phage prepared by the procedure mentioned above was infected to *E. coli*, for example to *E. coli* Y 1090(r), and proliferated on agarose plates to obtain cDNA library. Screening was carried out by using DNA of enhancer region of α -fetoprotein gene as a probe.

DNA coding for the protein which binds to the enhancer of α -fetoprotein gene could be obtained.

The following is a detailed description of the invention.

Example

This example shows the preparation of DNA coding for a protein that binds to the enhancer of α -fetoprotein gene from mRNA prepared from human hepatoma-derived cell line, HuH-7.

① Confirmation that the DNA probe having enhancer activity:

The fact that the enhancer activity resides on the DNA which was used as a screening probe was confirmed by the following CAT assay.

The DNA fragment having enhancer activity [(refer to Fig.1 and Molecular and Cellular Biology 8, 5179 (1988)], was isolated from the human genomic DNA by the digestion with restriction enzymes HgiA I and BstN I. The HgiA I/BstN I fragment was purified by agarose gel electrophoresis, and its termini were filled in with DNA polymerase I Klenow fragment. pUC18 plasmid was digested with BamH I and the ends were treated in the same manner. The two fragments were ligated using T4DNA ligase (Takara Shuzo Co., Ltd.) to give recombinant plasmid pAFE (HgiA I/BstN I)₁.

The plasmid pAFE(HgiA I /BstN I)₁ was subsequently digested with restriction enzymes Xba I and Pst I, and nucleotides were deleted from the ends with Exonuclease III and mung-bean nuclease (Takara shuzo). Then the ends were filled in with DNA polymerase I Klenow fragment and the plasmid was recircularized after ligating Xho I linkers to the ends.

E. coli DH5 was transformed with the recircularized plasmid and a clone which underwent deletion down to nine bases downstream from TTAATAAT sequence was isolated. This plasmid was designated as pAFE (HgiA I /Xho I)₁. The plasmid pAFE(HgiA I /Xho I)₁ was then hydrolyzed with restriction enzymes Nla IV and Xho I, and the fragment containing TTAATAA sequence was isolated by agarose gel electrophoresis. Ends of the Nla IV/Xho I fragment were filled in with DNA polymerase I Klenow fragment, and then the Nla IV/Xho I fragment was inserted into the Cla I site of plasmid pAF1.OCAT. The obtained plasmid was designated as plasmid pAF1.OE25CAT(see Fig.2).

The plasmid pAF1.0 CAT contains one kb DNA of the α -fetoprotein gene promoter region, structural gene for chloramphenicol acetyltransferase (CAT) of *E. coli*, SV40 poly(A) addition signal and t-antigen intron. The vector is used to screen for DNA fragment having enhancer activity [J. Biol.Chem., 262, 4812 (1987)]. Transfection of CAT plasmids into HuH-7 cells and CAT assay were performed as follows, based on the known method [Mol. Cell Biol., 2, 1044(1982)].

About 8×10^6 cells were grown in a 75 cm² culture flask. To these cells, 20 to 30 μ g CAT plasmid DNA was transfected by calcium phosphate method. The cells were incubated for two days in RPMI-1640 culture medium containing 0.3% lactalbumin hydrolysate, after which the cells were washed with PBS, and scraped off using a scraper. Then the cells were collected by low-speed centrifugation (800 rpm, 5min.), washed once with a solution composed of 40mM Tris-HCl(pH7.5), one mM EDTA and 150mM NaCl, and suspended in 100 μ l of 250 mM Tris-HCl (pH7.5). This suspension was repeatedly frozen and thawed five times to lyse the cells, and centrifuged in an Eppendorf centrifuge at 12,000rpm for five minutes to obtain a supernatant, which is then heat-treated at 65 °C for 10 minutes. The treated supernatant was again centrifuged at 12,000rpm for five minutes with the same machine to obtain a supernatant which was used as a cell extract.

After determining the protein concentration of the cell extract using the protein assay kit (BioRad), 50-200 µg of protein was used for the following CAT assay.

A typical CAT assay reaction mixture, 180 µl, contains 250mM Tris-HCl (pH 7.5), 0.1 µ Ci [¹⁴C]chloramphenicol (Amersham), 0.4mM acetyl-CoA, and 50-200 µg protein extracted from the cells. This reaction mixture was incubated at 37°C for 60-180 minutes, after which one ml of ethyl acetate was added to extract chloramphenicol and its acetylated derivatives. The ethylacetate layer was then separated and dried in vacuo. The residue was redissolved in 20 µl of ethyl acetate, a portion or all of which was applied onto a silica gel plate (Merk). The resultant plate was then developed using chloro-

form: methanol (96:4. V/V). The silica gel plate was dried at room temperature, wrapped with wrapping film, and subjected to autoradiography. As illustrated in Fig.3, a greater CAT activity was observed when the experiment was performed with CAT plasmid, pAF1.OE25CAT, which contained Nla IV/Xho I fragment derived from α-fetoprotein gene enhancer, than with pAF1.OCAT, indicating the presence of an enhancer activity in the Nla IV/Xho I fragment.

②Growth of α-fetoprotein producing cells:

Human hepatoma line, HuH-7, which produces α-fetoprotein was cultured under the following conditions. The cell line is available from Professor Jiro Sato, Pathology Division, Cancer Institute of Okayama University, Japan.

Conditions for culture:

The medium used was RPMI-1640 medium supplemented with 0.3% (W/W) lactalbumin hydrolysate (Gibco) or RPMI-1640 medium containing 5-10% (V/V) fetal calf serum. The medium was replaced as needed, and incubation was performed in an incubator (37°C) filled with air containing 5% carbon dioxide.

③Preparation of poly (A) RNA from the cells described above:

Total RNA was extracted from 2×10⁸ cells, adopting guanidinium isothiocyanate-caesium chloride method [Biochemistry, 18, 5294(1979)], by the following procedure.

20ml solution composed of sixM guanidinium isothiocyanate, five mM sodium citrate, 0.1M 2-mercaptoethanol and 0.5% sodium N-lauroyl sarcosinate was added to the cells, and the cells were homogenized at room temperature, then four g of cesium chloride was dissolved per 10ml of this homogenate. Over the cushion of 2.5 ml solution composed of 5.7 M cesium chloride and 0.1M EDTA (pH7.5), in a polyallomer tube, 10 ml of the homogenate was layered. The tube was then cen-

trifuged at 34,000rpm for 18 hours at 20°C in Hitachi Ultracentrifuge Rotor RPS 40T (Hitachi Ltd.). The resultant precipitate was dissolved in one ml solution composed of 10mM Tris-HCl (pH 7.4), 5 mM EDTA and 1% SDS. To this solution, an equal volume of a mixture of chloroform and n-butanol (4:1, V/V) was added, mixed well and the mixture was centrifuged (16,000g, 10min.). To the obtained aqueous phase, one-tenth volume of 3M sodium acetate (pH 5.5) and 2.5 volumes of ethanol were added and mixed well. The mixture was allowed to stand at -70°C for more than two hours and centrifuged (16,000g 20min.) to precipitate RNA. The precipitate was washed with 70% ethanol and then dried (total RNA).

For the preparation of poly(A) RNA from the total RNA, oligo(dT) cellulose chromatography was employed as described below. Oligo (dT) cellulose (50mg) was packed in a small column and equilibrated with a solution composed of 10 mM Tris-HCl (pH 7.5), 0.5M NaCl, one mM EDTA and 0.1% SDS. Then 390 µg of the total RNA dissolved in the same buffer was applied on the column. RNA that was not retained was washed out by the buffer, and the retained RNA was eluted with a solution composed of 10mM Tris-HCl(pH7.5), one mM EDTA and 0.05% SDS. To this eluate, 1/10 volume of 3M sodium acetate (pH5.5) and 2.5 volumes of ethanol were added, mixed, and the mixture was allowed to stand for more than two hours at -70 °C. The mixture was then centrifuged (12,000g 15min.). The precipitate formed was washed with 70% ethanol, dried and dissolved in sterilized distilled water. This fraction was designated as poly(A) RNA.

④Synthesis of cDNA:

cDNA synthesis kits purchased from Pharmacia and Amersham were used for this purpose.

(i) cDNA synthesis using random primer

To the first-strand synthesis reaction mixture (32 µl), 2.5 µg of poly(A) RNA and 0.4 µg of random primer, dp(N)6 (Takara Shuzo Co., Ltd.) were added and then allowed to react according to the supplier's protocol(Pharmacia). An EcoR I adaptor was ligated to both ends of double-strand cDNA as directed by the protocol. From 2.5 µg poly (A) RNA, 2.3 µg of doublestranded cDNA was obtained.

(ii) cDNA synthesis using oligo (dT) primer

One µg of double-strand cDNA was synthesized from 2.5 µg of poly(A)RNA according to the protocol provided by Amersham. EcoR I adaptor was ligated to the ends of double-strand cDNA as directed by the protocol. Finally, 200ng of double-strand DNA with adaptor was obtained.

⑤ Preparation of recombinant DNA and recombinant phages:

Protophone λ gt11 system (Promega) was used for the preparation of recombinant DNA. Ligation was performed according to the protocol. The recombinant DNA thus obtained was packaged into recombinant phage using Packer Gene Packaging System (Promega).

The process of preparing cDNA and λ gt11 recombinant phage is shown in Fig.4.

⑥ Preparation of λ gt11 cDNA library:

The recombinant phage packaged as described above was infected to *E. coli* Y1090(r) and λ gt11 cDNA library was prepared according to the known method [DNA cloning, Vol.1, p49, IRL Press (1985)].

⑦ Preparation of probes and the screening method:

The probe for the screening was prepared as follows from the DNA fragment of the human α -fetoprotein gene enhancer located upstream of the gene (see Fig. 4).

The DNA fragment having enhancer activity [described in Molecular and Cellular Biology 8, 5179 (1988)], was isolated by digesting genomic DNA with restriction enzymes *Hgi*A I and *Bst*N I. *Hgi*A I/*Bst*N I fragment was purified by agarose gel electrophoresis, and its termini were filled in with DNA polymerase I Klenow fragment. pUC18 plasmid was treated with *Bam*H I and the ends were similarly treated. The two fragments were ligated using T4 DNA ligase (Takara Shuzo Co., Ltd.) to give a recombinant plasmid pAFE (*Hgi*A I/*Bst*N I)¹.

The pAFE(*Hgi*A I/*Bst*N I)¹ was subsequently digested with restriction enzymes *Xba*I and *Pst*I and nucleotides were deleted from the ends with exonuclease III and mung-bean nuclease. Ends were filled in with DNA polymerase I Klenow fragment and the plasmid was recircularized by inserting *Xho*I linker to the ends.

A clone which underwent deletion down to nine bases downstream from TTAATAAT sequence was isolated from the deletion mutants obtained by transforming *E. coli* DH5 with the plasmid and designated as pAFE(*Hgi*A I/*Xho*I)¹.

The plasmid pAFE(*Hgi*A I/*Xho*I)¹ was then hydrolyzed with restriction enzymes *Nla*IV and *Xho*I, and a fragment containing the enhancer was isolated by agarose gel electrophoresis.

Ends of the *Nla*IV/*Xho*I fragment were filled in with DNA polymerase I Klenow fragment and the fragment was self-ligated using T4 DNA ligase. The catenated DNA was cloned into *Hinc*II digested pUC18 plasmid. *E. coli* DH5 α (BRL) was transformed with the recombinant plasmid. A clone containing six *Nla*IV/*Xho*I fragments ligated in tandem was selected from the recombinants and designated as pAFE(*Nla*IV/*Xho*I)₆.

This plasmid was cleaved with restriction enzymes

*Bam*H I and *Hind*III. 3' termini of the obtained *Bam*H I / *Hind*III fragment were labelled with DNA polymerase I Klenow fragment and 5'-[α -³²P]dCTP (Amersham) and used as a probe for screening.

⑧ Method of screening (1):

Screening of the λ gt11 cDNA library prepared from cDNA which was synthesized by using random primer, was performed according to the known procedure [Cell, 52, 415, (1988)]. As a result, cDNA clone λ 2 was isolated.

⑨ Method of screening (2):

Clones other than λ 2 were isolated by screening cDNA libraries with ³²P-labelled λ 2 cDNA according to the method of Benton, W.D. and Davis, R. W. [Science 196, 180 (1977)].

λ 130, λ 475, λ 476 and λ 488 were isolated by using λ 2 cDNA as a probe. λ 537 and λ 659 were isolated by using portions of λ 488 as probes. The positions of each cDNA clone are shown in Fig.5.

The fact that the protein encoded by the λ 2 cDNA, recognizes TTAATAATT sequence was confirmed by gel shift assay and footprint analysis, as described below.

⑩ Preparation of lysogen extracts:

E. coli Y1089 was transfected with the recombinant phage according to a known method [DNA cloning, Vol. 1, p.49; Ed. by D.M. Glover, IRL Press(1985)] to produce lysogen. Extract was prepared according to Singh et al. [Cell, 52, 415 (1988)]. As a control, extract was also prepared from *E. coli* Y1089 which was transfected with λ gt11.

⑪ Gel Shift Assay:

Gel shift assay was performed according to the known method [Nature, 319, 154 (1988)] using the above mentioned lysogen extracts and DNA probe prepared from pAFE(*Hgi*A I / *Bst*N I)¹ described in ⑦, as follows.

pAFE(*Hgi*A I/*Bst*N I)¹ was digested with *Sma*I and *Hinc*II to prepare the enhancer DNA fragment. Probe was prepared by labelling the 5'-ends of the *Sma*I / *Hinc*II fragment with ³²P.

As shown in Fig.6, two shifted bands (a and b) were detected with the extract made from λ 2 lysogen, while such bands were undetectable with λ gt11 lysogen extract. The result indicated that protein encoded by the cDNA cloned in λ 2 has a property of binding to the probe DNA.

⑫ Identification of protein-binding site on the DNA fragment by footprint analysis:

λ2 lysogen extract containing 16 µg protein (as total protein) and Sma I/Hinc II probe, prepared from pAFE (HqIA I/BstNI 1) described in ⑦, were mixed in a 75 µl of reaction solution containing 10mM of Tris-HCl (pH 7.5), 50mM NaCl, one mM DTT, 0.5mM of EDTA, 2.5mM of MgCl₂, three µg poly (dl-dC) · poly(dl-dC) and 5% glycerol, and allowed to stand at room temperature for 30 minutes. The probe used in footprint analysis is labelled with ³²P on only one of the 5' termini. To this mixture, one µl of DNase I (40 ng/µl) was added and the mixture was incubated for one minute at room temperature, after which 2.3 µl of 0.5M EDTA was added to cease the reaction. The reaction mixture was applied to a 5% polyacrylamide gel, and was subjected to electrophoresis at 11 V/cm until bromphenol blue migrated close to the lower end of the gel.

After removing one glass plate, the gel was wrapped with Saranwrap (Asahi Chemical Ind. Co., Ltd.) and autoradiographed.

The portion of gel corresponding to the position of DNA bound to protein and that of unbound DNA, determined by the relative mobility of the bands on the autoradiogram, was cut out of the gel. The gel section was immersed in 0.5ml solution composed of 0.5 M ammonium acetate, 0.1% of SDS and one mM of EDTA, and allowed to stand overnight at room temperature to extract the DNA. The mixture was centrifuged (12,000g, 10 min.) to obtain the supernatant. The gel section was washed twice with 0.5ml each of the same elution buffer and the wash was combined with the eluate. To this eluate, five µg of yeast transfer RNA (Sigma) was added and the resultant mixture was extracted with a mixture of phenol and chloroform (1:1). To the aqueous phase thus obtained, two volumes of ethanol was added and then the mixture was allowed to stand for 30 minutes or longer at -70°C. The DNA was then collected by centrifugation (12,000g, 10min.), washed once with 70% ethanol and was dried under vacuum. This DNA was dissolved in 10 µl solution composed of 80% formamide, 10mM of NaOH, one mM of EDTA, 0.1% xlenecyanol and 0.1% bromophenol blue, heated at 90°C for three minutes, and 3-4 µl of the solution was applied to an 8% polyacrylamide gel containing seven M urea, which is commonly used to determine nucleotide sequences. The samples were subjected to electrophoresis. Size markers prepared from the probe DNA by chemically cleaving at guanine and adenine residues [Methods in Enzymology, 65, 499 (1980)] were run side by side with samples.

The result of autoradiography performed after electrophoresis are shown in Fig.7. As shown in the figure, it was indicated that, in DNA bound to the protein encoded by λ2 cDNA, the 15 base-pair portion containing TTAATAATTA in the middle was protected from cleavage by DNase I.

⑬ Nucleotide sequence of the cDNA

Phage DNA was prepared according to the conventional method [Molecular Cloning, a Laboratory Manual, p.76 (1982)]. After hydrolysis with EcoR I, insert DNA fragment was separated by agarose gel electrophoresis. The purified insert DNA fragment was subcloned in plasmid pUC118 or plasmid pBluescript II SK⁺ and amplified in *E. coli* DH5 α.

Fig.8 shows the nucleotide sequence of the cDNA, determined by dideoxy method [Method in Enzymology, 65, 560 (1980)] and Maxam-Gilbert method [Method in Enzymology, 65, 499 (1989)], and the amino acid sequence of the protein encoded by the cDNA.

Several nucleotide sequence heterogeneities were found in overlapping clones. At nucleotide 2316, λ 537 has cytosine (C), but λ488 has thymine (T).

At nucleotide 4338, λ659 has thymine (T), but λ 488 has cytosine (C). At nucleotide 5565, λ130 has thymine (T), but λ488 has cytosine (C). At nucleotide 7508, λ475 has cytosine (C), but λ476 has thymine (T). At nucleotide 8403, λ475 has guanine (G), but λ476 has thymine (T). Among these, the mutation at nucleotide 7508 resulted in the replaced amino acid residue (alanine residue in λ475 versus valine residue in λ476). Differences derived from deletion and insertion were also found in addition to the point mutations. An adenine (A) at nucleotide 4378 was deleted in λ488, however it was not deleted in λ659. Deletion of three consecutive nucleotides at two places, namely from nucleotide 7903 to 7905 and from 7921 to 7923, and insertion of 24 nucleotides (GG CAA CTA CAG CAG CAG CAG CAG C) between nucleotide 7528 and 7529 were observed in λ476 in comparison with λ475. These alterations resulted in an increase of six amino acid residues in the polypeptide encoded by λ476.

The nucleotide sequence shown in Fig.8 is based on λ475.

DNA encoding the full length enhancer binding protein was constructed by ligating the cDNAs described above and inserted to pBluescript II SK⁺. *E. Coli* harbouring the recombinant plasmids were deposited at the Fermentation Research Institute.

pBSAFP1A/DH5α (deposit number BP-3147) has the nucleotide sequence of λ476. pBSAFP1B/DH5α (deposit number BP-3148), on the other hand, has the nucleotide sequence of λ475. Both pBSAFP1A/DH5α and pBSAFP1B/DH5α have thymine, cytosine and cytosine at nucleotide 2316, 4338 and 5565, respectively.

Claims

Claims for the following Contracting States : AT, BE, CH, LI, DE, DK, FR, GB, IT, LU, NL, SE

1. A DNA which codes for a protein that has an amino

acid sequence given in Fig. 8 and that binds specifically to the human α -fetoprotein gene enhancer.

2. A DNA which has a nucleotide sequence given in Fig. 8, in which a cytosine at nucleotide 7508 may be replaced by other nucleotides and, as a consequence, the amino acid 2460 (alanine) in the encoded protein may be replaced by other amino acids.
3. A DNA described in claim 2 in which a cytosine at nucleotide 7508 is replaced by a thymine.

Claims for the following Contracting State : ES

1. A DNA which codes for a protein that has an amino acid sequence given in Fig. 8 and that binds specifically to the human α -fetoprotein gene enhancer.
2. A DNA which has a nucleotide sequence given in Fig. 8, in which a cytosine at nucleotide 7508 may be replaced by other nucleotides and, as a consequence, the amino acid 2460 (alanine) in the encoded protein may be replaced by other amino acids.
3. A DNA described in claim 2 in which a cytosine at nucleotide 7508 is replaced by a thymine.
4. A method of producing a biologically active protein comprising inserting a DNA which codes for a protein that has an amino acid sequence given in Fig. 8 and that binds specifically to the human α -fetoprotein gene enhancer into an expression vector and co-transfecting animal cells with another expression vector in which a gene for the biologically active protein is placed under the control of α -fetoprotein gene enhancer and promoter.
5. A method according to claim 4 in which the cytosine at nucleotide 7508 may be replaced by other nucleotides, preferably thymine, and the amino acid 2460 (alanine) in the encoded protein is replaced by other amino acids.

Patentansprüche

Patentansprüche für folgende Vertragsstaaten : AT, BE, CH, LI, DE, DK, FR, GB, IT, LU, NL, SE

1. DNA, die ein Protein kodiert, das eine in Figur 8 angegebene Aminosäuresequenz aufweist und das spezifisch an den menschlichen α -Fetoprotein-Genverstärker bindet.
2. DNA, die eine in Figur 8 angegebene Nukleotidsequenz aufweist, worin ein Cytosin am Nukleotid

7508 durch andere Nukleotide ersetzt werden kann und als Konsequenz die Aminosäure 2460 (Alanin) in dem kodierten Protein durch andere Aminosäuren ersetzt werden kann.

3. DNA nach Anspruch 2, worin ein Cytosin am Nukleotid 7508 durch ein Thymin ersetzt wird.

Patentansprüche für folgenden Vertragsstaat : ES

1. DNA, die ein Protein kodiert, das eine in Figur 8 angegebene Aminosäuresequenz aufweist und das spezifisch an den menschlichen α -Fetoprotein-Genverstärker bindet.
2. DNA, die eine in Figur 8 angegebene Nukleotidsequenz aufweist, worin ein Cytosin am Nukleotid 7508 durch andere Nukleotide ersetzt werden kann und als Konsequenz die Aminosäure 2460 (Alanin) in dem kodierten Protein durch andere Aminosäuren ersetzt werden kann.
3. DNA nach Anspruch 2, worin ein Cytosin am Nukleotid 7508 durch ein Thymin ersetzt wird.
4. Verfahren zur Herstellung eines biologisch aktiven Proteins, das das Einfügen einer DNA, die ein Protein kodiert, welches eine in Figur 8 angegebene Aminosäuresequenz aufweist und spezifisch an den menschlichen α -Fetoprotein-Genverstärker bindet, in einen Expressionsvektor und das Co-Transfizieren von tierischen Zellen mit einem anderen Expressionsvektor umfaßt, in dem ein Gen für das biologisch aktive Protein unter die Kontrolle eines α -Fetoprotein-Genverstärkers und Promotors gebracht wird, umfaßt.
5. Verfahren nach Anspruch 4, worin das Cytosin am Nukleotid 7508 durch andere Nukleotide, vorzugsweise Thymin, ersetzt werden kann und die Aminosäure 2460 (Alanin) in dem kodierten Protein durch andere Aminosäuren ersetzt wird.

Revendications

Revendications pour les Etats contractants suivants : AT, BE, CH, LI, DE, DK, FR, GB, IT, LU, NL, SE

1. ADN qui code pour une protéine qui a la séquence d'acides aminés donnée dans la Figure 8, et qui se lie spécifiquement à l'amplificateur du gène de l' α -fétoprotéine humaine.
2. ADN qui a la séquence de nucléotides donnée dans

la Figure 8, dans lequel une cytosine au nucléotide 7508 peut être remplacée par d'autres nucléotides et, en conséquence, l'acide aminé 2460 (alanine) dans la protéine encodée peut être remplacé par d'autres acides aminés.

5

3. ADN décrit dans la revendication 2, dans lequel une cytosine au nucléotide 7508 est remplacée par une thymine.

10

Revendications pour l'Etat contractant suivant : ES

1. ADN qui code pour une protéine qui a la séquence d'acides aminés donnée dans la Figure 8, et qui se lie spécifiquement à l'amplificateur du gène de l' α -fétoprotéine humaine.

15

2. ADN qui a la séquence nucléotidique donnée dans la Figure 8, dans lequel une cytosine au nucléotide 7508 peut être remplacée par d'autres nucléotides et, en conséquence, l'acide aminé 2460 (alanine) dans la protéine encodée peut être remplacé par d'autres acides aminés.

20

3. ADN décrit dans la revendication 2, dans lequel une cytosine au nucléotide 7508 est remplacée par une thymidine.

25

4. Procédé de production d'une protéine biologiquement active, dans lequel on insère un ADN qui code pour une protéine qui a la séquence d'acides aminés donnée dans la Figure 8, et qui se lie spécifiquement à l'amplificateur du gène de l' α -fétoprotéine humaine dans un vecteur d'expression, et on cotransfecte des cellules animales avec un autre vecteur d'expression dans lequel un gène pour la protéine biologiquement active est mis sous la commande de l'amplificateur et du promoteur du gène de l' α -fétoprotéine.

30

35

40

5. Procédé selon la revendication 4, dans lequel la cytosine au nucléotide 7508 peut être remplacée par d'autres nucléotides, de préférence la thymine, et l'acide aminé 2460 (alanine) dans la protéine encodée, est remplacé par d'autres acides aminés.

45

50

55

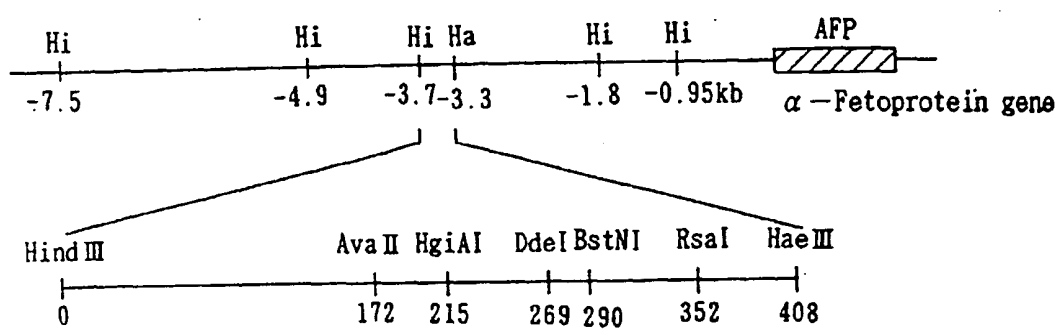


Fig. 1

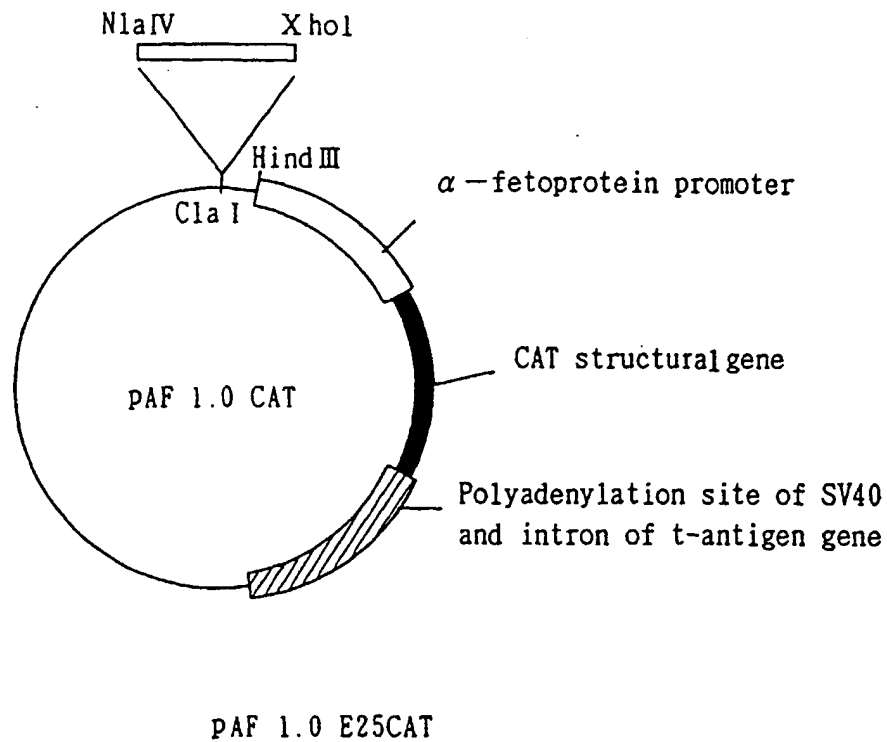


Fig. 2

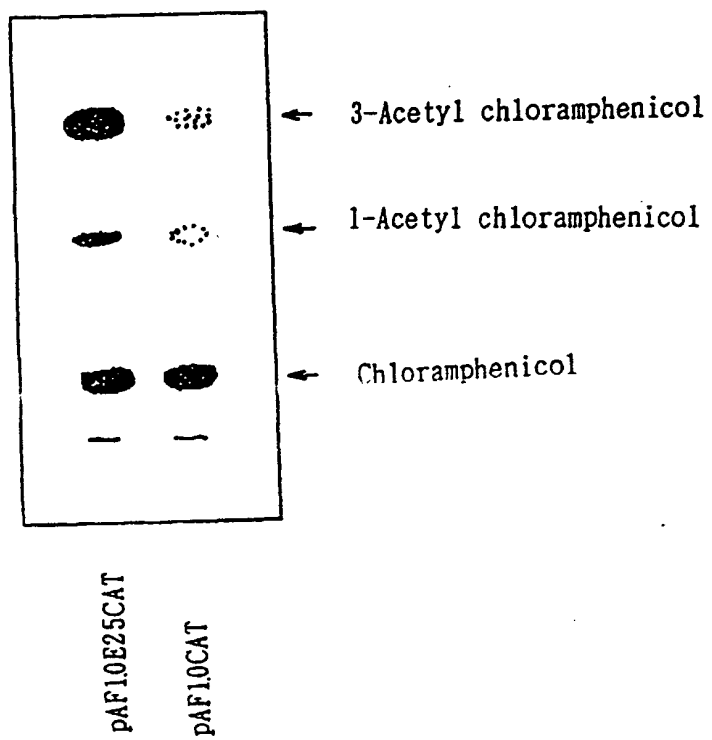


Fig. 3

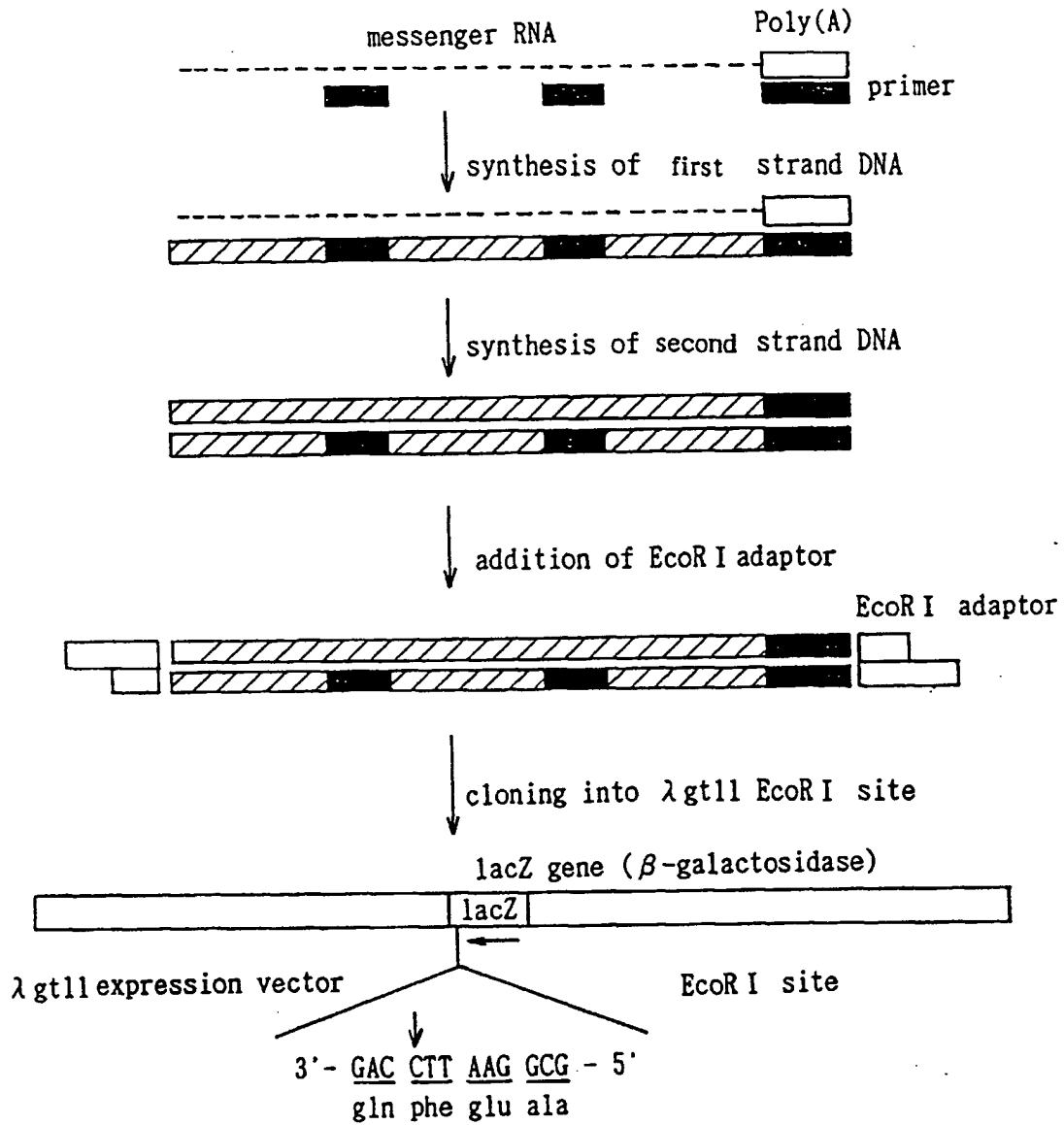


Fig. 4

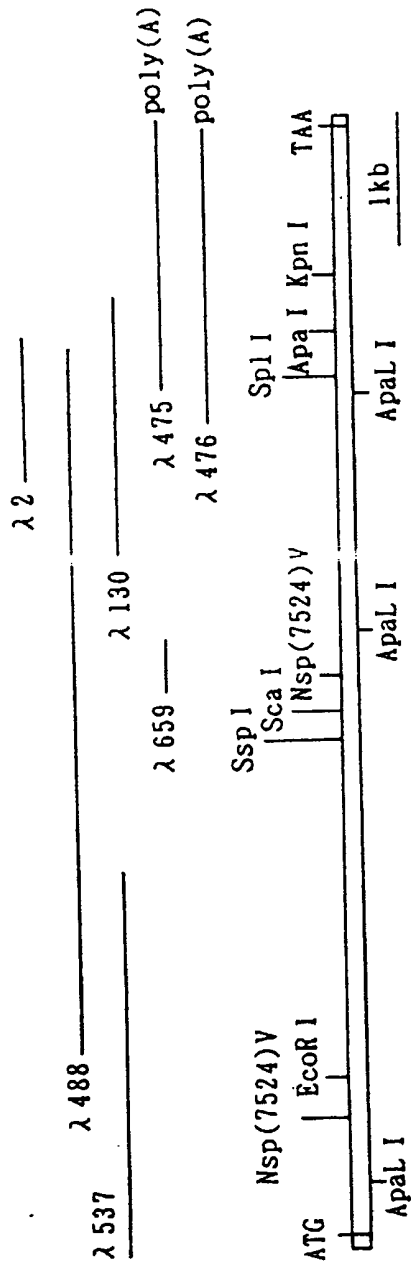


Fig. 5

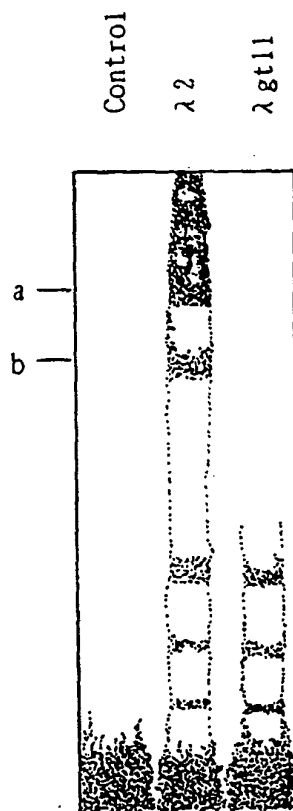


Fig. 6

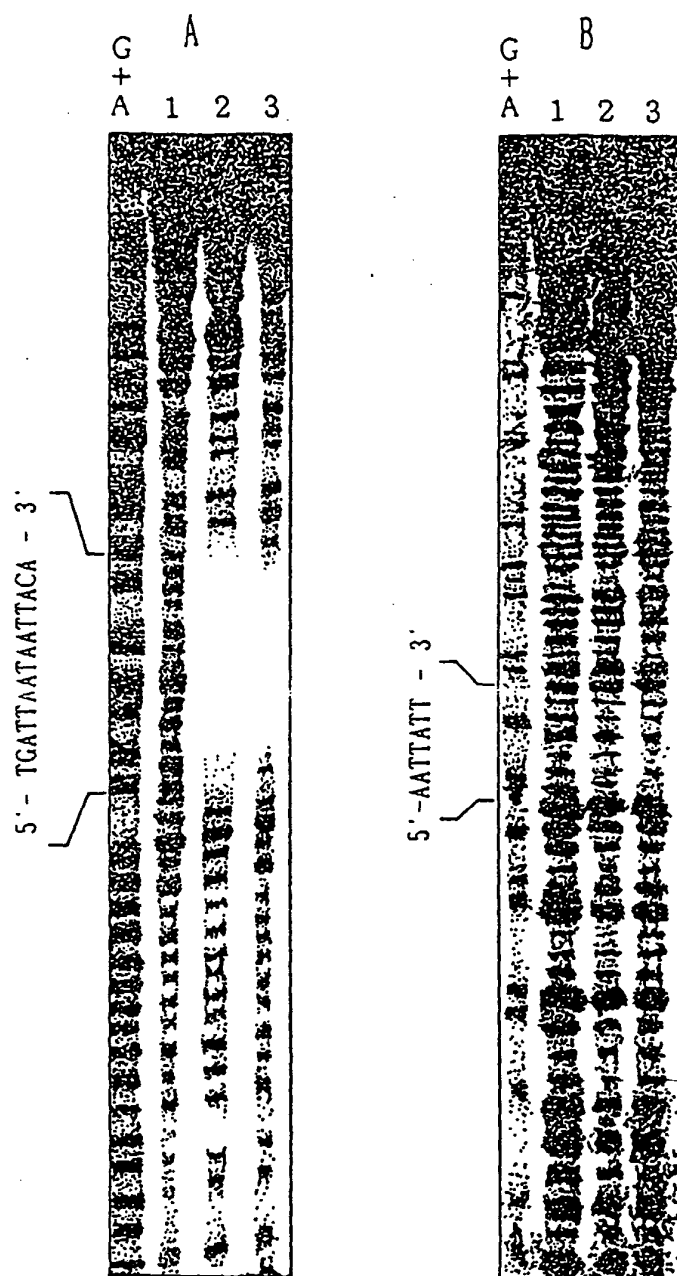


Fig. 7-a&b

Fig. 7-c

17

AACAAAGATCCACCTCGACCTGACCCACCTCCACAGCGTGGACCTGACCTGAGAGGCTCATATGACGGTGACCC 1260
 N N K I H L Q L H L T H L H S V A P D C V E X L I M T V T T
 CCTGAGATGGTGCAGCAGCAGCATGTTCTCCAGGAGCTGTCCAGATGGAGATGGGAATTCATTTGGAAGCGCAGGAAGCAG 1350
 P E N V N P S S H F L P A A V P D R D G N S N L E E A G K Q
 CCTGAAACCTCAGAGGATCTGGGAAAGAACATCTTGGCATCCGCAAGCAGCAGAGCAAGCGGAGATTGAAACCATCCCTGCTGACCCA 1440
 P E T S E D L G K N I L P S A S T E Q S G D L K P S P A D P
 GGCCTGTGAGAGAGACTCAGGCTTCATCTGCTGGAGAGGGGTGCAACGAGTTTTCAAAACCTCTGCTGCCCTTCAGACGCATTTT 1530
 G S V R E D S G F I C V K K G C N Q V F K T S A A L Q T H F
 AATGAACTGCATGCCAAGGCCCTCAGCTGCGGCTGTGAGATCGGCAATGTACAAAGTACCGCTGTATCAGTGTAGCTGGCCCTTCAG 1620
 N E Y H A K R P Q L P V S D R H V Y K Y R C N Q C S L A F K
 ACCATTGAAAAGTTGCAGCTCCATTTCTCAGTACCATGCTGCAAGCTGCCACCATGTGCTGTCTTTGTGACGGCAGTTTCGGAATTC 1710
 T I E K L Q L H S Q Y H V J R A A T M C C L C Q R S F R T F
 CAGGCTCTGAAGAGCCACTTGAGCAAGCCACTGGAGCTGAGTGAGGCTGACATCCACAGCTTTATGGTGGCTGGCCCAATGGG 1800
 Q A L K K H L E T S H L E L S E A D I Q Q L Y G G L L A N G
 GACCTCTGGCAATGGAGACCCCACTCTGGCTGAGGACCAATACCATAATTTGAGGAGACAAAGGAGAGAGTGTGACTTGGAGAT 1890
 D L L A N G D P T L A E D H T I I V E E D K E E S D L E D
 AACAGAGCCCAACGGGCACTGCTGGGTGAGTACAGAGAGACTCGGCTCAGAGCCCAAGAGAGCTCTGCCCTTTCAGAAAAGGTCCC 1980
 K Q S P T G S D S G S V Q E D S G S E P X R A L P F R K G P
 AATTTACTATGGAAAAGTTCTAGACCCCTTCTCGCCCTTACAAAGTGTACCGCTGCGAAGGAATCTTTCACCTCAAAAAGAAATATCTGCTA 2070
 N F T N E K F L D P S R P Y K C T V C K E S F T Q K N I L L
 GTACACTACAATTTCTGCTCCCACTGCATAGTTAAAGAGAGCCCTTCAAGAAATCAGCAACCGGTACGCCAGAACCCACAGCAGCCCA 2160
 V H Y N S Y S H L K L K R A L Q E S A T G Q P E P T S S P
 GACAAACAACCTTTAAGTGTAACTGTGATGTGGCTACAGGAGAGTTCCACTCTGAGATCCATATGAGGTCTGTGTATCATCA 2250
 D N K P F K C N T C N V A Y S Q S S T L E I H M R S V L N Q
 ACCAAGGCCCGGAGCCCAAGCTGGAGGCTGCAAGTGCCAGCAGCAATGGGACTGGGAAACAGCAGCAATTTGCTTGAGCTCGTCAGG 2340
 T X A E A A K L E A A S G S S N G T G N S S I S L S S T

Fig. 8-c

TTCAAACAGCTCGAGAGGTTTGCCAAACAGTACAGAGACCACTACGATAAACTGTACCCACTGAGGCCCCAGACCCCGAGGCCACCA 3510
 F X Q L E R F A K Q Y R D H Y D K L Y P L R P Q T P E P P P
 CCTCCCCCTCCACCCCTCCACCCCACTTCGGGCGAGCGCGCTCAGCGGGCTCCACACCAAGCCATCCCGCATCAGGCCCAACCAATC 3600
 P P P P P P P P L P A A P P Q P A S T P A I P A S A P P I
 ACCTCACCTACAAATTGCACCGGCCCAAGCATCAGTGCGGCTCAGCCAGCTCTCCATGCCATGGAGCTGCCCATCTTCTCGCCGCTGATG 3690
 T S P Y I A P A Q P S V P L T Q L S N P N E L P I F S P L K
 ATGCAGACGATCCCGCTGCAGACCTTGCAGGCTCAGCTACCCCGCAGCTGGGACCTGTGGAGCCCTCAGCCCTCGGACCTGGCCCAATC 3780
 H Q T H P L Q T L P A Q L P P Q L G P V E P L P A D L A Q L
 TACCAGCATCAGCTCAATCCAAACCTCTGCTCCAGCAGCAGAAACAGAGGCTCGCACCCAGGATCAGATGATCAGCTCCGAGTCTTCCGG 3870
 Y Q H Q L N P T L L Q Q Q N K R P R T R I T D Q L R Y L R
 CAATATTTGACATTAACAACCTCCCCAGTGAAGAGCAAAATAAAGAGATGGCAGACAAAGTCCCGGTTGCCCCAGAAAGTGATCAAGCCAC 3960
 Q Y F D I N N S P S E E Q I K E H A D K S G L P Q K V I K H
 TGGTTCAGGAACACTCTCTTCAAAGAGAGCGGCTAACAAAGGACTCCGCTTACAACCTTCAAGTAACTCCTATCACCAGCCGAGGAG 4050
 V F R N T L F K E R Q R N K D S P Y N F S N P P I T S L E E
 CTCAGATTGACTCCCGCCCTTCGCGGAACCTCCAAAGCAGGAGTACTGGGGAAGCAGAGAGGCTTTCAGAAACAAAGGTTTACGGAC 4140
 L K I D S R P P S P E P P K Q E Y V G S K R S S R T R F T D
 TACCAGCTGAGGCTCTTACAGGACTTCTTCGATGCCAATGCTTACCAAGGATGATGAATTTGAGCAACTCTCTAATTTACTGAACCTT 4230
 Y Q I R Y L Q D F F D A N A Y P K D D E P E Q L S N L L N L
 CCAACCCGAGTGTAGTGGTGTGTTTCAGATCCCGACAGAAAGCCAGGAAGAATTTATGAGAAATCAGGGAGAGGGCAAGATGGAGAG 4320
 P T R V I Y V V F Q N A R Q K A R K N Y E N Q G E G K D G E

11: 09 00 1 0

4410	CGCGTGAGCTTACAAATGATAGATACATTGGAACAAGCAACTTGAACCTACCAAGTGCACAAAATGTACCTGTGTTTCAGCGCATCTTT R R E L T N D R Y I R T S N L N Y Q C K K C S L V F Q R I F
4500	GATCTCATCAAGCACCAGAGAAGCTGTGTTACAAGGATGAGGATGAGGAGGCGAGGACGACAGCCAAAATGAGDATTCCATGGATGCC D L I K H Q K K L C Y K D E D E G Q D D S Q N E D S H D A
4590	ATGGAATCTGAGCGCTACCAAGCTCATCTGCACTACCCGATACCCGATTCACAGGCTTACAGGCCGCCACGACACCATCAGGCCAATAATACA N E I L T P T S S C S T P H P S Q A Y S A P A P S A N N T
4680	GCTTCCTCCGCTTCTTGACGCTTACAGCGGAGGCTGAGGAATCTGGCAACCTTCAATTCAAAAACAGAGGCGCGGATGAGAAACCAAAG A S S A F L Q L T A E A E L A T F K S K T E A G D E K P K
4770	CTGGCGGAAGCTCCCACTGCACAGCCAAACCAACCAAGCAAGGACAAACCAAGGCGAGAGCTGCAGCAGCAAGAGCAAGCCCGAG L A E A P S A Q P N Q T Q E K Q G Q P K P E L Q Q E Q P E
4860	CAGAAAGCAACACTCCACAGCAGAGCTCCGCCAGCTGGTGTCCTGCTTGCACAGGCTCCTCCACAGGCGGCCCTCCACAG Q K T N T P Q Q K L P Q L V S L P S L P Q P P Q A P P P Q
4950	TGCCCTTACCECAGTCGAGCGCCAGTCTTCCAGCTCTCCAGCTCTCCAGCCCTGCAAGCCCTCCACACATCAACTCTCTCAACAGCTCGCA C P L P Q S S P S Q L S H L P L K P L H T S T P Q Q L A
5040	AACTTACCTCTCAGCTAATCCCCTACCAGTGTGACCAAGTGAAGTGGCAATTCGGTCAATTTGAGCACTGGCAGGAGCAATCAGCAGCTC N L P P Q L I P Y Q C D Q C K L A F P S F E H V Q E H Q Q L
5130	CACCTTCTGAGCGCGCAGAACCAAGTTCCACCCCGAGTTTTTGGACAGGTCCCTGGATATGCTTTTCATGCTCTTTGATCCAGTAAAC H F L S A Q N Q F I H P Q F L D R S L D H P F H L F D P S N
5220	CCACTCTGGCCAGCCAGCTGCTCTCTGGGGCCATACCTCAGATTCAGCAAGCTCAGCCACTTCTCTTCACTCCAAACCTCCACAATG P L L A S Q L L S G A I P Q I P A S S A T S P S T P T S T H
5310	AAACACTCTCAAGAGGAGCTGGAGGAAAAGGCCAGTCCAAAGCCCTGGCGAAAACGACAGTGGGACAGGAGAGAGCCTCAGAGAGAC N T L X R K L E E K A S A S P G E N D S G T G G E E P Q R D
5400	AAAGCTTTGAGAACCAACCATCACACCGGAAACAACCTAGAAAATCTCTACCAAGAACTATCTACTGGATTCCAATCCGACTCGAAAAGATGTTG K R L R T T I T P E Q L E I L Y Q K Y L L D S N P T R K N L

5490 GATCACAATGCCAGGCTTGAAGAAACGCTGTGGTACAAAGCTGGTTTCAGAAACACCCGAGCTCGGAAAGGAAAGGACAGTTTC
 D H I A H E V G L K K R V V Q V V F Q N T R A R E R K G Q F
 5580 CGGGCTGTAGCCCGCAGGCCACAGGAGATGCCCTTTTTCAGAGCGCTCTTCAAAGCCAAAGACTGCTCTTGAGGCTCATATCCGG
 R A V G P A Q A H R R C P F C R A L F K A K T A L E A H I R
 5670 TCCGTCAGTGGCATGAAGCCAAAGAGAGCTGGCTACAACTAATCTGTCTCGGATGCTCTTACACTGTGTAGTGGGGGACTCCAGATGAAA
 S R H V H E A K R A G Y N L T L S A H L L D C D G G L Q H K
 5760 GGAGATATTTTGGCGGAACCTAGCTTTTCCCACTACCCCAAGCAAGTGTGAGTGGTTCAGGGTGTCCCGCTCTCACTGTGAGTAAACCC
 G D I F D G T S F S H L P P S S S D G Q G V P L S P V S K T
 5850 ATGGAATTGTCAACCAAGAACTCTTCTAAGCCCTTCCCTCCATTAAGGTGGAAGGATTAAGAGACTTTGAAGCCCTCCATGTCTCAGTT
 H E L S P R T L L S P S S I K V E G I E D F E S P S N S S V
 5940 AATCTAAACTTTGACCAAACTAAGCTGGACAAACGATGACTGTTCCTCTGTCAACACAGCAATCACAGATACCACTGGAGACGAGGC
 N L N F D Q T K L D N D D C S S V N T A I T D T T T G D E G
 6030 AACCGAGATAACGACAGTCCAAACGGGAATAGCAACTGAAGCAAAATCTCTTCTGCACCCCAAGCAAGGGTTGACCAAAAGCGCCATGATG
 N A D N D S A T G I A T E T K S S S A P N E G L T K A A H M
 6120 GCAATGTGAGTATGAAGATCGGTTGTCTGCTGGTCAAGCCCGCCCGAGCTTTTATACCAAGGAATATGACAATGAAGGTACA
 A H S E Y E D R L S S G L V S P A P S F Y S K E Y D N E G T
 6210 GTGGACTACAGTGAACCTCAAGCCCTTGCAGATCCCTGCTCCCGAGTCCCTGGTCCGAGTGGATCTGCAGGCAAAATCTGGTGACACGCGG
 V D Y S E T S S L A D P C S P S P G A S G S A G K S G D S G
 6300 GATCGGCTGGCCAGAAAGGTTTTCGCACTCAAAATGACCAATCTGCAGCTGAAGGTCCTCAAGTCAATGCTTTAATGACTACAGGACACCC
 D R P G Q K R F R T Q N T N L Q L K V L K S C F N D Y R T P
 6390 ACTATGCTAAGTGTGAGGTCCTGGGCAATGACATTTGGACTGCCAAAGAGAGTCTGTTCAAGTCTGGTTCAGAGAAATGCCCGGCAAAAGAA
 T H L E C E V L G N D I G L P K R V V Q V V F Q N A R A K E
 6480 AAGAAGTCCAAAGTAAAGCATGGCAAGCAATTTTGGTATAAAACCAACGAGTTATGAGGGAACCAAAACAGAGTGCACCTTTGTGTGGCATC
 K K S K L S H A K H F G I N Q T S Y E G P K T E C T L C G I
 6570 AAGTACAGCGCTGGCTGTCTGTACGTGACCATATCTTTCCCAACAGCATATCTCCAAAGTTAAGACACCATTTGGAAGCCAGCTGGAC
 K Y S A R L S V R D H I F S Q Q H I S K V K D T I G S Q L D
 6660 AAGGAGAAAGAAATACCTTTGACCCAGCCACCGTAGCTACGTTGATGGCTCAACAAAGAGTTGGACCGGATTAATAAGGCAACGAGGTCTT
 K E K E Y F D P A T V R Q L H A Q Q E L D R I K A N E V L

Fig. 8-9

GAGTGTCCCCCTCTCCGAAATCCCTGAAGAGCCAGAGCCAGAAAGTCCGGACTCCCTCTACGACCCCTTCATTGTICCA 7740
 E V S P L L P K L P E E P E A E S K S A D S L Y D P F I V P
 AAGTGCAGTACAAAGTTGGTCTGCCGCAAGTCCAGGGCGGCTTCAGCGACAGAGGAGCCAGAGGAGCCACCTGAAGTCCCTCTGCTTC 7830
 K V Q Y K L V C R K C Q A G F S D E A A R S H L K S L C F
 TTCGGCCAGTCTGTGCTGTAACCTGCAAGAGATGGTCTTCACGTCCCAACCGCGCGCGGTGGCAGTGGCGCGCGCGCGCGGT 7920
 F G Q S V V N L Q E H V L H V P T G G G G G G S G G G G G G
 GCGCGCGCGCGCGCGCGCGCGCTCGTACCACTGCTGCGGTGCGAGAGCGCGCTCTGTGGGAGGAAAGCTCTGAGTCAACAT 8010
 G G G G G G G G G G G G S Y H C L A C E S A L C G E A L S Q H
 CTCGAGTCGGCCCTTGCACAAACACAGAAACAATCACGAGAGCAGCAAGAAACGCCAAAGAGCACCTAGTTTATTACCTCACCTGCTGC 8100
 L E S A L H K H R T I T R A A R N A K E H P S L L P H S A C
 TTCGCGATCCTAGCACCGCATCTACCTGCGAGTCTGCCGCTCACTCAAAACGACAGCCCGCTCCCGCGTCCCGCGCGCGCGCTCC 8190
 F P D P S T A S T S Q S A A H S N D S P P P P S A A A P S S
 GCTTCCCCCAGCGCTCCAGGAAGTCTTGGCCGCAAGTGTCTCCGGGCTTCGGCAGCGAAGCCCGCTCTTTTCTCTCTCTCTCA 8280
 A S P H A S R K S W P Q V V S R A S A A K P P S F P L S S
 TCTTCAACGGTTACCTCAAGTTCATGCAACACCTCAGGGGTTCAAGCCCTCGATGCCAAGCAGACGACTATTCGGAGGAGTCTGACACGGAT 8370
 S S T V T S S S C S T S Q V Q P S H P T D D Y S E S D T D
 CTCAGCCAAAGTCCGACGGACCGCGGAGCCCGGTGGAGGGTCCCAAGAGACCCCGCTGCCCAAGCAGACAGTGGTCTGACCAAGTACGA 8460
 L S Q K S D G P A S P V E G P K D P S C P K D S G L T S V G
 ACGACACCTTCAGATTGTAAAGCTTTGAAGATGAACAATACAAACAATGAATTTAATAACAAAAATTAACAACCAATTTICAAAAA 8550
 T D T F R L
 TAGACTAACTGCAATTCCAAAGCTTCTAACCAAAAAACAAAAA
 8601

Fig. 8-b